STATISTICS SEMINAR

Shengnan Chen

Preliminary Exam

Monday, November 8, 2021

Dickens 109, 2:30 pm (Central Time)

Poisson Regression with Laplace Measurement Error

The count data appear in many areas including biostatistics and econometrics. Poisson linear regression has been a popular model for fitting such data. The intensity function of the Poisson regression is often modeled as a linear function of the covariates. Statistical inferences on the Poisson linear regression are abundant in literature. However, often times the covariate cannot be measured precisely due to potential measurement instruments imperfection or human errors, and we can only observe a surrogate of the unobserved covariate. That is, the covariate is contaminated with measurement error. Many estimation procedures for Poisson regression are proposed in literature when the error follows normal distribution. However, there has been few works about statistical inferences when the errors follow Laplace distribution, while Laplace errors are quite common in the real practices. In this dissertation, novel estimation procedures are proposed for a class of Poisson linear regression when the covariate is contaminated with Laplace measurement error.

In the first project, we propose a weighted least squares estimation procedure that incorporates the first two conditional moments of the response variable given the observed surrogate, and the weight function is intentionally chosen to avoid the complexity caused by the random denominator and to increase the estimation efficiency. To solve for the conditional moments, a Tweedie-type formula has been adopted. Instead of assuming the distribution of the unobserved covariate is known, we assume that the distribution of that latent variable is unknown. Large sample properties of the proposed estimator, including the consistency and the asymptotic normality, are discussed. The finite sample performance of the proposed estimation procedure is evaluated by simulation studies, showing that the proposed estimator is more efficient than the existing ones.

In the second project, we propose a corrected maximum likelihood estimation procedure based upon a Tweedie-type formula for the conditional expectation of the likelihood function given the observed surrogate. Two situations, the distribution of the latent variable is known as well as unknown, are considered. Large sample properties of the proposed estimator are discussed, and simulation study shows that the estimator is more efficient than the existing estimation procedures.

Although the discussion is conducted for univariate cases, the proposed estimation procedure can be readily extended to the multivariate cases by using multivariate Tweedie-type formulae.